COA

Chapter 04: |Cache Memory

Prof. Dr. Mostafa Elhosseini

https://youtube.com/drmelhosseini

AGENDA

- C/Cs of memory systems
  - Location
  - Capacity
  - Unit of transfer
  - Access method
- Memory hierarchy

Intro…

- Computer memory is organized into a hierarchy
- As one goes down the memory hierarchy, one finds decreasing cost/bit, increasing capacity, and slower access time
- The **design challenge** is to organize the data and programs in memory so that the **accessed memory words** are usually in the **faster memory**

Characteristic of Memory Systems

Location

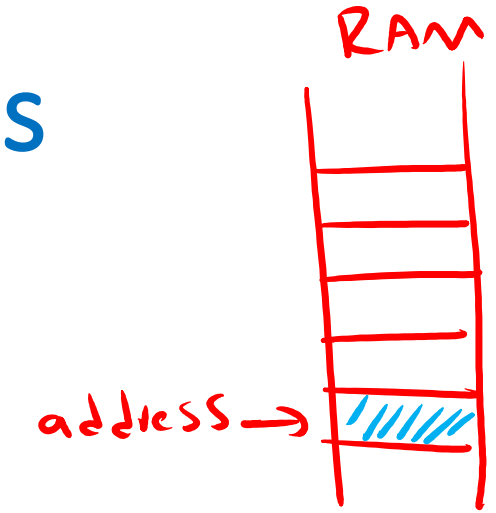
- whether memory is **internal** and **external**
  - Registers
  - Cache
  - Main Memory
- **External memory** consists of peripheral storage devices, such as disk and tape, that are **accessible to the processor via I/O controllers**

Characteristic of Memory Systems

Capacity

- For **internal memory**, this is typically expressed in terms of **bytes** (1 byte 8 bits) or **words**.
- **Common** word lengths are **8, 16, and 32** bits.
- **External** memory capacity is typically expressed in terms of **bytes**.

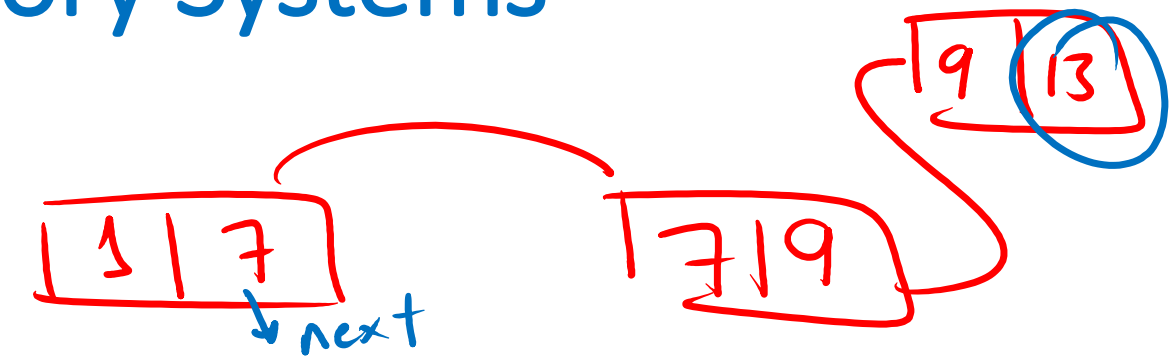
Characteristic of Memory Systems



Unit of Transfer

- For **internal** memory
  - Data Bus Width
  - May be equal to the word length, but is often larger, such as 64, 128, or 256 bits
- For **external memory**, data are often transferred in much larger units than a word, and these are referred to as **blocks**
- **Addressable units**
  - Smallest location which can be uniquely addressed
  - Word – Byte

Characteristic of Memory Systems



Access Method

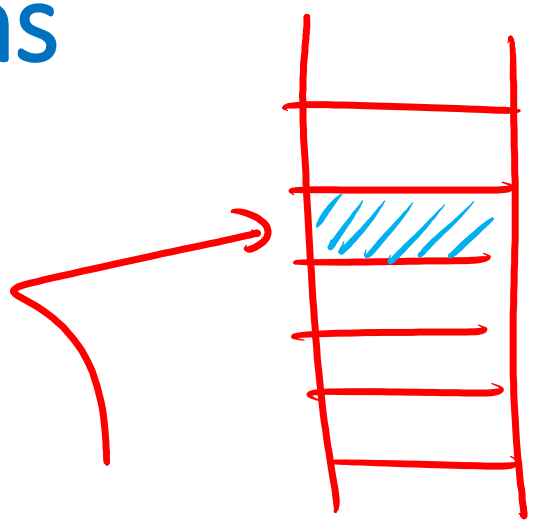
- Sequential access
  - Start at the beginning and read through in order
  - Access time depends on location of data and previous location
  - Stored as a **linked list**, in which each piece of data contains a reference to the following piece
  - Looking for information in a book with no index
  - e.g., audio or video or cassette tape

Characteristic of Memory Systems

Access Method

- Direct access
  - Individual blocks have unique address
  - Access is by jumping to vicinity plus sequential search
  - Access time depends on location and previous location
  - e.g., disk

Characteristic of Memory Systems



Access Method

- Random access
  - Individual addresses identify locations exactly
  - any location can be selected at random and directly addressed and accessed
  - Access time is independent of location or previous access
  - e.g., RAM – some caches

Characteristic of Memory Systems

Access Method

- Associative
  - As with ordinary random-access memory, each location has its own addressing mechanism, and retrieval time is constant independent of location or prior access patterns
  - Data is located by a comparison with contents of a portion of the store
  - word is retrieved based on a portion of its contents rather than its address.
  - Access time is independent of location or previous access
  - e.g., cache

Characteristic of Memory Systems

Performance

- Access time
  - The time it takes to perform a read or write operation
- Memory cycle time
  - Access time plus any additional time required before a second access can commence.
  - This additional time may be required for transients to die out on signal lines or to regenerate data if they are read destructively
- Transfer Rate
  - Rate at which data can be transferred into or out of a memory unit
  - Equal to 1/(cycle time).

Memory Hierarchy

- There is a **trade-off** among the three key characteristics of memory: namely, capacity, access time, and cost
- Not to rely on a single memory component or technology, but to employ a memory hierarchy